NASA

World Wide Web
Best Practices

Prepared By:

NASA Web Best Practices Team
Chartered By The NASA Chief Information Officer and the Principal Center for Communications Architecture

Last Update: September 1, 1998

Table of Contents

Web Best Practices Team Members

Justin Jackson/MSFC
Brian Dunbar/HQ
Jeff Delabeaujardiere/GSFC
Dabney Hibbert/HQ
Emma Antunes/GSFC
Robin Land/LARC
Rick Weller/JSC
Steve Eubanks/LERC
Randy Sparkman/MSFC
Gaye Graves/ARC
Sue Wheaton/DFRC

Executive Summary

The Internet and the World Wide Web have become an integral component of the provision of NASA information to academia, industry, and the public-at-large, and as an enabling set of technologies for the internal NASA information technology environment. This document provides a set of web "Best Practices" (WBP) that are intended to serve as a guideline to all NASA entities engaged in the development and maintenance of World Wide Web (WWW) resources. The guidelines defined in this document are applicable to all NASA Enterprises, Centers, and project and program offices and have been created under the auspices of the NASA Chief Information Officer (CIO).

The Best Practices as defined in this document are organized into the following sections:

〈 Site Design
〈 Structure/Site Organization
〈 Navigation
〈 Page Design
〈 Site Accountability
〈 Security
〈 Accessibility
〈 Marketing
〈 Searching
〈 Authoring Tools

It is the intent of this report to facilitate consistent, reliable, and efficient deployment of WWW services throughout the Agency. These criteria should be viewed as baseline recommendations. References to applicable government and NASA policies and standards are included to highlight "requirements." This document will continue to be enhanced and extended by the NASA WWW community.

1.0 Introduction

The provision of information to NASA partners, academia, industry and the public-at-large is a fundamental component of the NASA charter. The Internet and the World-Wide-Web have become an integral element of this task. Further, these same technologies also serve as a core piece of the internal information technology (IT) infrastructure on which NASA relies to carry out its mission. In this context, the set of technologies, content, processes, and policies that support the delivery of WWW-based services should be viewed as a significant part of the NASA environment and approached with the same commitment to excellence as all other NASA activities.

2.0 Purpose

This document provides a set of "Best Practices" that are intended to serve as a guideline to all NASA entities engaged in the development and maintenance of WWW resources. It is the intent of this report to facilitate consistent, reliable, and efficient deployment of WWW services throughout the Agency. These criteria should be viewed as baseline recommendations. References to applicable government and NASA policies and standards are included to highlight "requirements." This document will continue to be enhanced and extended by the NASA WWW community.

The current document and associated information is available at:

http://nasa-wbp.larc.nasa.gov/

3.0 Scope

The guidelines defined in this document are applicable to all NASA Enterprises, Centers, and project and program offices. It has been created and will be maintained under the auspices of the NASA Chief Information Officer (CIO).

Due to evolving technology and changing government policy, this guidelines document will be reviewed by a NASA agency team on a four month cycle. The last review date is displayed on the cover of the document. Documents with review dates exceeding four months may contain obsolete or inaccurate information, and a later version should be obtained from the NASA WWW Best Practice site at (http://nasa-wbp.larc.nasa.gov/).

4.0 Site Design

The two most important factors in web design are:

〈 To know the main goals and objectives in order to effectively communicate the purpose of the site.
〈 To identify the target audience in order to understand the motivations, goals, and behaviors of that target audience.

Developing a clear sense of purpose and articulating this to the intended audience, either explicitly or implicitly, is vital to meeting expectations when presenting the content of the web site. Careful attention to primary goals and objectives will lend focus and clarity to the content of the site and its organization. Indeed, some of the references indicate that a clear goal statement should be included in the content -- possibly even an outline of the site's contents.

For example, what priority does a site give to publishing recent developments, such as the NASA Home Page (http://www.nasa.gov) versus another providing access to catalogs of information, such as the National Space Science Data Center (NSSDC) Web site (http://nssdc.gsfc.nasa.gov/). Giving high priority to recent developments will drive other decisions, including allocating space within the design for new material and the need for easy updates. If the material isn't going to change very often, but encompasses extensive data holdings, then design and maintenance issues may be less of a concern than site structure and network capacity.

In order to communicate effectively with the intended audience, one must understand their motivations for visiting a site and their behavior in navigating and using the information. A clear understanding of how the target audience will perceive a site will help in determining what content, level of detail, format, and navigational aids are employed. An Intranet site, for example, should target a captive audience with well-organized content, which will allow frequent users to quickly and accurately obtain their desired information. A public site, however, needs to provide context and easy access while drawing the user further into its content. Many groups have spent large amounts of time trying to determine who NASA's audience is. The truth is the Agency has a very broad, but also segmented, audience. Most often it has been divided into these kinds of groups:

〈 **General Public and Mass media:** These audiences typically have some general knowledge.
〈 **Research communities beyond NASA; industry; scientific and technical press:** This audience typically possesses in-depth knowledge about specific subjects.
〈 **Government groups (Congress, the Administration, other agencies):** may be knowledgeable or not.
〈 **Internal:** May be focused on a Center, a project, or even a small organizational group.

〈 **Academia:** Teachers, administrators and students. (Recently this group has become a much higher priority for both the Federal Government Administration and NASA. However, it has also become almost a catchall for doing anything from full-production videos to Web sites. NASA employees who are honestly interested in reaching these groups must bear in mind they need very specific kinds of material tied to curricula. Employees should work with their Centers' education groups to ensure that Web sites meet these needs.)

References:
〈 Yale Style Manual -
http://info.med.yale.edu/caim/manual/sites/site_elements.html
〈 WWW Federal Consortium Guidelines and Best Practices-
http://hoohoo.ncsa.uiuc.edu:7777/consortium_info/agency_policy_guidelines/agency_policy.html

4.1 Structure/Site Organization

Keeping in mind high-level goals and the perspectives of the target audience, the structure of most web sites is typically hierarchical to some extent. This is almost exclusively the case with Intranet sites, where structure lends navigability to a typically dense information content. Though often true for public content as well, a hierarchical structure is often amended when necessary with content designed to draw the visitor into the site. Hierarchical structures, or tree structures should be well balanced; neither too deep to avoid losing the user, nor too shallow— presenting the user with too many options. The hierarchy should represent some logical division of the overall content, keeping important information toward the top. A metric of structural organization is the number of mouse clicks required for a user to locate information. A good set of steps to keep in mind:

〈 divide pages/information/content into logical units,

〈 establish a hierarchy of importance and generality,

〈 use the hierarchy to structure relationships among chunks

〈 for practicality and efficiency, analyze the functional and aesthetic aspects of the site.

- Excerpted from the *Yale Style Manual*
(http://info.med.yale.edu/caim/manual/sites/site_design.html)

Also of value is determining the relative importance of browsing and navigating— both as a content author, as well as from the user perspective. If, after analysis, it is determined that navigation is more significant, there may be of value providing some form of overview, possibly employing site maps to assist the user in quickly locating the content they seek. If the emphasis of a site is to create an "experience" for the user (more of the

browsing model), the overhead in producing and maintaining site overviews or maps may not be warranted.

References:

⟨ Yale Style Manual -
http://info.med.yale.edu/caim/manual/sites/site_design.html

⟨ NCSA TRG Review of Web Styleguides -
http://www.ncsa.uiuc.edu/edu/trg/styleguide/index-r.html#3

4.2 Navigation

A well-designed site should simplify the user's ability to locate information of interest, and facilitate the user's ability to move from one page to another. Part of this can be addressed in the structure of the site, but part must also be addressed by providing the user a means of identifying his or her current location in the site's information space (and time) as well as, providing tools to help the user move from one location to another. The following general rules of thumb should be applied to aid the user in navigating through a site's content.

Provide some visual context
⟨ Visual techniques such as employing common document templates, style sheets, and even such mundane items such as last updated/modified date stamps or displaying the URL of the page will provide clues for the user to understand where/when they are within the organization of information.

Minimize reliance on browser navigational buttons
⟨ Wherever possible, work navigational links directly into web content so that the progression from one document to another is intuitive for viewers who might otherwise be unfamiliar with site-specific button-type navigational aids, or for users who might have entered at a point which did not provide context for those tools.
⟨ If buttons are used as navigational aids, make sure the buttons are as self-explanatory as possible, and avoid the temptation to create navigational aids which distract the user's attention away from the site's actual content.

Provide a Search mechanism
⟨ Past a certain level of complexity, it typically becomes impossible to fully address the navigability of a site purely through structure and navigational aids. At this point, it is typically recommended that a site provide some sort of search interface or site index to permit visitors and users to more directly locate the information they are seeking.

References:
- ⟨ WWW Federal Consortium's Guidelines and Best Practices - http://hoohoo.ncsa.uiuc.edu:7777/consortium_info/agency_policy_guidelines/agency_policy.html
- ⟨ NCSA TRG Review of Web Styleguides - http://www.ncsa.uiuc.edu/edu/trg/styleguide/index-r.html#6
- ⟨ Yale Style Manual - http://info.med.yale.edu/caim/manual/intro/purpose.html
- ⟨ Sun Styleguide - http://www.sun.com/styleguide/others/Navigation.html

4.3 Content (Overall)

File formats
- ⟨ The file formats employed on a web site should be examined from both a macro-site-perspective as well as from a page-design perspective. The primary criterion for selection of file formats employed site-wide should be whether a specific file format can be easily used by the target audience of that content. For publicly accessible sites, file formats should be selected which make minimal assumptions regarding the ability of a visitor to handle formats not intrinsically supported by prevalent web browsers. This in and of itself is not an easy task since the number of file formats directly handled by most web browsers is constantly growing. If in order to increase the usability and/or appeal of a public site, non-natively supported file formats are used, care should be exercised to employ formats for which "readers" can be easily obtained (typically downloadable from the Internet at no cost), and have a significant installed base.
- ⟨ In closed-target communities such as Intranets where there is greater control, and predictability exists over the tools/facilities available at the desktop; additional file formats may be adopted based on commonly available applications (e.g., MS-Word, Adobe portable document format (PDF)). However, be aware of the Center's baseline software suite. Many centers discourage user-installed applications outside those tested and approved by the Center's information technology group. It is very frustrating to users to find a NASA site designed with plug-ins or other software that the users do not have and are supposed to obtain on their own.

References:
- ⟨ WWW Federal Consortium Guidelines and Best Practices - http://hoohoo.ncsa.uiuc.edu:7777/consortium_info/agency_policy_guidelines/agency_policy.html

4.4 Maintainability

4.4.1 The Art of Re-Use

It makes practical sense while developing the context of the site (including hypertext markup language (HTML), images, and other objects) to develop the context in such a way that it can be re-used in order to reduce total site maintenance. This can be done in several ways:

⟨ Use shorter, more readable HTML documents - this will aid in locating problems and making corrections as the content evolves.

⟨ Store frequently used content components like: images, headers, footers, etc. in shared, centralized locations rather than replicating the same content in multiple locations on the server. This reduces duplicate maintenance activities for similar components throughout the site.

⟨ Use global templates/style sheets or server-side includes to reduce the number of files that need to be maintained.

⟨ Avoid using browser-dependent features that require having two or more versions of HTML code, and maintaining differing versions.

4.4.2 Links

Links within a web site are probably the single largest maintenance issue and often the biggest single source of frustration for visitors. Effort should be made to minimize the likelihood a visitor will experience an invalid or broken link by:

⟨ Carefully balancing the number of links offered with the quality of those links. The more links employed raises the probabilities that someone will move a target document and break a link.

⟨ Don't move documents or change uniform resource locators (URL's) too often. It invites the possibility of broken links or breaking bookmarks that visitors have made to a site.

⟨ If the site must be moved then either post a pointer page or redirect the user to the new site from the old URL. A pointer page notifies the user the site has moved to a new URL, the location of the URL, and a direct link to the new page location. Generally pointer pages only need to remain at the old site for approximately 3-6 months. A redirect is configured within the web server to push the user to the new site while bypassing the old site altogether. This alleviates a stop at the old site, as required by a pointer page. However, sometimes users do not realize they have been redirected to another URL, and they will not bookmark or modify their site link to the new location.

⟨ Using relative URLs between pages within a site will simplify the

11

administrative aspect of relocating the content, should that be necessary.

〈 Use some form of automation wherever possible to identify those broken links that will inevitably occur.

4.4.3 Automation

In any site of even nominal size where even due care has been exercised in content development, maintenance quickly exceeds the limitations of the effectiveness of manual procedure. The only way to effectively combat this problem is to put in place various forms of automation to assist in identifying problems, and in some cases provide content creation.

〈 Link validation - Rather than attempting to manually scan hundreds and even thousands of documents for broken links, one should employ link validation utilities, and run these utilities at regular intervals to identify and report broken links.

〈 HTML validation - Employ these tools to identify HTML coding errors or HTML features which may not be supported by all browsers. Often these tools are run before finally releasing a document on a server, but these tools too could be periodically run on regularly updated pages to ensure their accuracy.

〈 Where possible, use automation to generate content based on select inputs minimizing the potential for human error. Obviously, sites which generate their content automatically from databases or other data sources eliminate or significantly reduce the amount of page and site maintenance-- only the information stored in the database then need be maintained. Keep in mind that it can be fairly expensive to have a database site set up. This would not be cost effective for small sites or sites that don't require frequent changes.

4.4.4 Feedback

Despite even the best attempts to ensure that links are accurate, errors will occur. Make sure that visitors can send feedback so that they can help identify problems.

References:

〈 The Sev Guide to Web Design: Maintenance - http://www.sev.com.au/webzone/design/maintenance.htm

〈 Web Site File Organization - http://www.glover.com/ss/extra08.html

〈 NCSA - TRG: Review of Web Styleguides -

〈   Art and the Zen of Web Sites - http://www.tlc-systems.com/webtips.shtml

〈   See the "Guidance on Implementation of NASA Website Privacy Statement" in the appendix.

4.4.5 Variable data

To avoid the constant upkeep of a page, avoid adding data that will constantly change. Rather than state "the software was developed five years ago", which will have to be changed every year, state "the software was developed in 1995." Rather than say "the current version of the software is 18.0," try to either avoid mention of a variable like that, or state "based on the 6/98 version of the software, version 18.0."

4.5 Architecture

4.5.1 Physical location of server

Like all site design issues, the physical location of the web server is significantly determined by the target audience. It is also significantly determined by the security needs and policies of an organization as well as the complexity and dynamics of the site and content.

As web servers add their own set of security risks to any IT environment, web sites exclusively targeting internal customers in an Intranet environment should be placed inside a firewall to reduce the risk of having them hacked from outside the organization.

Similarly, sites exclusively targeting external audiences should be placed (once appropriate procedures have been taken to secure the server as much as possible), outside of the firewall to not only facilitate communications with the customers it serves, but to also reduce the likelihood that the web server could be used to hack into a facility's local networks.

Typically a site will need to serve both internal and external customers. Numerous possible solutions and architectures exist to address this issue, often involving multiple web servers, proxy servers, and firewall configuration issues. The degree of compartmentalization and specifics of implementation must take into account

appropriate security needs and policies, the complexity and dynamics of the web content itself, and the ease of use— both from a user and content-developer standpoint.

References:
〈 <u>Web Security, A Step-by-Step Reference Guide</u>, by Lincoln D. Stein, Addison-Wesley, 1998.
〈 World Wide Web Security Frequently Asked Questions (FAQ) - http://www.w3.org/Security/Faq/

4.5.2 NASA Integrated Information Technology Architecture

〈 The NASA Integrated Information Technology Architecture (NIITA) document [Current rev: NASA-STD-2814] is an evolving document designed to establish a common set of goals and guidelines that provide a blueprint for NASA IT service providers, and as such defines a common, vendor independent framework for design, integration and implementation of IT systems. In this document it is indicated that "NASA has ... embraced the use of the World-Wide-Web (WWW) Browser as a 'universal' client and can be considered a leader in the use of WWW clients and servers for the dissemination and retrieval of information". The NIITA document also indicates that "NASA has thoughtfully considered the idea of an 'Intranet' and what it means for the Agency and the industry-at-large. The general consensus within the Agency is that an Intranet is the structured use of the open, scaleable technologies of the Internet (such as those listed in the preceding paragraph) to do the work of an Enterprise in a secure and reliable manner".
  〈 Though the NIITA does not provide specific guidance with respect to web site design, it does define a set of basic principles which would generally apply:
  〈 All IT system design must be customer driven.
  〈 IT solutions should demonstrate a commitment to a standards-based, modular design.
  〈 IT solutions should utilize stable, commercially available solutions if at all possible.
  〈 IT solutions should leverage existing IT investment.
  〈 Design objectives should give preference to general-use solutions.
  〈 Design objectives should give preference to a simpler IT infrastructure that anticipates multiple, and emerging, end-user devices.

References:
〈 NASA-STD-2814 - http://www.hq.nasa.gov/office/cio/standards/2814.doc

4.5.3 Browser Plug-ins

A list of NASA approved plug-ins is under development. This list, which will be coordinated by the Web Best Practices Team, Lewis Research Center as the Principal Center for Desktop and Workgroup Software, and the NASA WebMasters, will be published at a later date.

5.0 Page Design

5.1 Logical vs. Physical Markup

A well-designed page should simplify the user's ability to locate information of interest, and facilitate the user's ability to move from one page to another. This can, in part, be addressed in the layout of the page, but part must also be addressed by providing the user adequate means of navigation on the page.

Use logical markup in favor of physical markup.

HTML includes tags for both *logical* and *physical* markup. Logical markup specifies the nature of a piece of text: this is a heading, a paragraph, an important word, a book title. Physical markup specifies how the text will look on screen: large font, bold face, italics. It is better to use logical markup whenever possible instead of physical markup and to let the browser choose how to display the specified text, because the desired appearance (e.g., bold) may not be available in a particular browser (e.g., on a character terminal or text-to-speech reader). In addition, the logical elements are generally more meaningful for later parsing of the document by humans or machines (declaring that something is a level 1 heading is more useful than saying it is in a bold, courier, 24pt font).
The number of logical markup tags is presently limited, but XML (Extensible Markup Language) will allow authors to invent their own tags and, in combination with style sheets, specify their preferred appearance.

> *Design tip: You may someday need to dissect the structure of all your pages by machine, and this will be easier if the logical structure is clear.*

5.2 Creating a Consistent Look

It is advisable to maintain a consistent appearance and navigation style among the pages on a web site. This helps users identify the site, recognize when they have left it, and navigate it successfully. Elements that are typically chosen to be similar on each page are the structure, colors, backgrounds, headers, footers, sidebars, and navigation elements. Maintaining consistency is easier with advance planning, server-side includes, templates, scripts, and style sheets.

> *Design tip: You may view your site as starting from the home page, but many users will also find you through a search engine and arrive at some intermediate*

page within your site. Such users should be able to identify your site and recognize that subsequent pages are part of it.

5.2.1 Structure

Use logical divisions consistently.

Each document typically comprises headings, text paragraphs, lists, tables, and other such logical divisions. Using these in a similar manner on all pages wherever possible is the first step towards maintaining a uniform look.

5.2.2 Colors

Use black text on default or solid white background.

Each browser has a default set of colors for the page background and the text, typically gray or white for the background, black for the text, blue for hypertext links not yet visited, and purple for visited links. Some browsers allow the page author to modify these colors, and it is tempting to give a document or site a unique look by doing so. However, it is difficult to find color combinations that are readable on all monitors, and changing the visited/unvisited colors may confuse viewers. If the background color is changed, also specify the text colors to avoid conflict with the reader's default colors.

5.2.3 Backgrounds

Avoid background images.

Some browsers allow the page author to specify an image to be shown as a page background. Typically the image is tiled to fill the window area. Background images should be avoided because they add download time to the page and slow the layout. They can make the page unreadable when text overlays parts of the image. Consistently aligning them with the text or other elements is impossible, and they often look poor when the page is viewed in a very large or very small window. If a background is used, the image size should be as small as possible and the image should be subtly varying and not high-contrast. A possible exception to the latter is the case of a margin along one edge of the page, perhaps in conjunction with a sidebar. However, in this case, as the window size is changed the alignment of the sidebar text may be destroyed and, for wide windows, the margin may be repeated in the middle of the page.

5.2.4 Headers

Use a thin, fast-loading banner (or no banner).

In this context a *header* refers not to the HTML headings tags (`<H1`, `<H2`, and so on), but to text and/or graphics that appear at the top of a page. If headers are included, use a similar one for each document. Examples of headers might be a line of text, a logo, or a banner that is short and wide. Banners that are different, but similar in concept, can be used for each subsection of the site. The header can even incorporate a path summary for site navigation, as at http://www.useit.com/. Because it takes up valuable screen space, the header should not be too many pixels tall. A single line of text wrapped in a table is one option. Any image used should be designed to load quickly as it is one of the first things in the document.

5.2.5 Footers

Use the footer for navigation elements and for site-wide information too long or uninteresting to fit in the header.

A footer is text and/or graphics that appears at the bottom of the page. As with the headers, footers should be similar throughout the web site. Because the footer appears at the end of the document, it can safely be taller than the header and can easily be used for navigation icons or menus. Footers might also include site or Agency logos, contact addresses, and feedback requests.

5.2.6 Sidebars

Use sidebars judiciously, favoring headers and footers when possible.

An alternative to both headers and footers is a sidebar, which is a vertical column of text along the left (typically) edge of the page. The advantage of a sidebar is that, unlike a header, it does not push the main contents farther down the page and, unlike a footer, it is visible when the page is first loaded. A sidebar allows prominent display of a large number of links without pushing other content completely off the first screen. The disadvantage is that it reduces the width available for each line of text, which is particularly a problem on small monitors or for users who choose not to use their browser full-screen. Also, text-only or speech-enabled browsers will tend to display the sidebar before any of the content, particularly if the sidebar is in the left margin.
Note that if tables are used to create a sidebar, the page may load less quickly because most browsers do not render anything until the entire table has arrived. This problem can be addressed by making the sidebar itself a table but putting the rest of the content in a normal page body. If frames are used, consider the WBP recommendations for their judicious use.

5.2.7 Navigation Elements

On every page, provide links to the major sections of the site.

Navigation elements refer to text or image links that allow the user to navigate among the major sections of a site. For example, if there is a row of icons in the footer, or a menu in a sidebar, or a path summary in the header, at a minimum, every page should have a link to the home page or to a higher-level page in the same section. Better yet, include links to the entry points for the important divisions of a site. Center or Agency rules may require links farther up the Agency hierarchy.

To maintain a unified look for a site and to avoid confusion for the user, choose a consistent set of navigation elements and repeat them on every page.

> *Design tip: A user arriving at your site from a search engine may start at some obscure page in the middle of the site structure and will need links back up to the major sections.*
>
> *Design tip: Given Section 4.4.2 Links, you should add a statement about giving links meaningful names. Do not simply create a link that reads "Click here to get information on subject X" or that reads "click here." Links should be self-explanatory where possible and should be part of the normal sentence structure of the site.*

5.2.8 Writing for the Web

Although the web provides a boundless amount of information, most people find reading from the screen an uncomfortable experience. Scrolling down lengthy pages becomes tedious. As a result, most users tend to skim or "surf" the page, reading titles, hypertext links, bulleted lists, and short paragraphs. As an alternative, people may print out pages to read later off-line.

To write content that will be read by the impatient web surfer, it's best to:

〈   Include your most important information in the lead sentence of short two or three sentence paragraphs.

〈   Content should be brief and it should be no longer than one and a half screens.

〈   Use hypertext links to highlight important information and lead the reader through your content.

〈   Since people may enter a website on any page, your information should not be dependent on concepts presented on "higher level" pages. However, using well-planned links may avoid the necessity of presenting redundant information.

〈 Titles for sections of homepages should be short, no more than one or two words, if they are to be used on buttons for navigating the website.

The Sun Microsystems *Guide to Web Style* is an excellent reference for developing written content for the web. It can be found at http://www.sun.com/styleguide/

5.3 Maintaining a Consistent Look

Once a consistent look has been established for a site, it is necessary to maintain it as pages are added, modified and deleted. The task is simplified by advance planning, server-side includes, templates, scripts, and style sheets.

5.3.1 Advance Planning

When starting or revising a site, spend some time planning it before writing HTML. Consider the major topics and subtopics, the structure and layout of a typical page, and the navigation among pages. Settling such issues in advance will avoid the need to redo pages. Advance planning may even mean developing a hierarchical chart of the site. It is useful in maintaining consistency of the levels, and when the site is complete, the chart can become the site map.

5.3.2 Server-Side Includes

Server-side includes (SSIs) refer to a feature of some servers that allows insertion of one document into another before the server ships the page over the network. For example, a footer, which is the same on every page, can be stored in a separate file and inserted by the server. Modify the footer document and all pages that refer to it are immediately updated. SSIs, if available on the server, are an easy way to help enforce consistency. Their disadvantage is that response is slightly slower because it takes time to parse a document and insert the requested file(s). Also, some servers do not set a last-modified time (for the benefit of caches) on server-parsed documents.

5.3.3 Templates

A template is a more general example of the SSI technique. If there are a large number of similar pages on the site, each with some static and some variable content, a template can be used to describe the static portions of the page. Tokens placed within the template are examined by the server and replaced by variable contents as needed. For example, a page showing information about a person might have static headers and navigation icons and variable tokens representing name, address and phone number. The server might fetch those variables from a database and insert them before shipping the page.
Using templates requires either custom programming or a third-party product.

5.3.4 Scripts

Another way to maintain uniformity is to generate pages dynamically using scripts. If a collection of similar pages where each one can be described using a small number of arguments (for instance, an image catalog where each page is uniquely identified by the image it displays), then a server-side script can determine what page the user wants and then create it on demand. Every page looks similar because the same script makes each page, and a site-wide change can be implemented simply by making changes to that script. Using scripts typically requires custom programming and places extra load on the server. However, scripts are the only way to handle the nearly infinite permutations of content extracted from a database.

5.3.5 Style Sheets

Style sheets are emerging as the preferred way to describe and maintain the appearance of pages on a site. A style sheet indicates what different elements on the page look like. A stylesheet can be coded in the html page, or it can be a separate file that is linked into the page. If you have a large site with many pages, it is best to have a separate file. This allows you to change the overall style of a web site by only changing one file, rather than changing each of the individual html pages. Each page can refer to that style sheet and thereby ensure that every time a particular element (say, a level 1 heading) is used its appearance is consistent.
As of mid-1998, the latest browsers do not fully support style sheets and are not consistent in their implementation. Therefore, it is important to research the elements of style intended for a site and ensure that these are compatible on all browser versions and platforms. Good places to research this are (**Error! Bookmark not defined.**), (**Error! Bookmark not defined.**) or (**Error! Bookmark not defined.**)
 Style sheet handling should improve with time, and the result will be a powerful tool.

5.4 Graphics

Graphics are used in web pages as navigation elements, logos, illustrations, decoration, and layout elements. Considerations for each of these vary, but there are general guidelines to follow:

⟨ Keep file size as small as possible.

⟨ Provide useful ALT attributes. (ALT is an attribute in an HTML image tag that provides a meaningful text alternative when images are not loaded.)

〈 Provide an alternate means of navigation that does not require images.

〈 Remember that search engines cannot read images.

〈 Never use images from other sites without permission.

〈 Specify WIDTH and HEIGHT attributes to allow the browser to render the document more quickly.

〈 Consult, and if appropriate link to, NASA's "Reproduction Guidelines for Use of NASA Images and Emblems" (http://www.NASA.gov/gallery/photo/guideline.html)

5.4.1 Navigation Elements

Navigation elements include both link icons and image maps. Icons are an easy way to give a site a unique but consistent look. Icons can both illustrate and link to the major sections of the site. To avoid confusion, a particular icon should always provide the same function.

Make the icons small (in both pixels and bytes) so they load quickly and do not overwhelm the page content.

The difficult part about icon creation is making their meaning obvious. A small picture cannot always convey an abstract topic (e.g., consider inventing an icon for the topic "alkalinity.")

Libraries of icons exist that can be downloaded and used. These are simple to acquire, but do not make a site unique. Consider hiring an artist to design appropriate icons.

5.4.2 Logos

A discreet logo on each page helps unify a site.

An Agency, Center, or project logo is a good way to identify a site and its sponsors. Make it easy to load and unobtrusive, as visitors are more interested in content than in a logo. Choose an alternative text that describes or replaces the logo ("Laboratory for Atmospheres" rather than just "logo").

Institutional requirements may dictate that the logo of the Agency or Center be displayed; consult the appropriate Public Affairs Office.

5.4.3 Illustrations

Choose and produce illustrations with care.

Images are essential for illustrations and figures in a document. Compared to a black-and-white academic journal with page charges, or to an expensive glossy brochure, the web

offers a cheap way to overload a document with large color graphics. This temptation should be resisted; instead, choose images wisely, using the same care as for an expensive print publication. Keep the file size as small as practical--crop or shrink the image, reduce the number of colors, use Joint Photographic Experts Group (JPEG) instead of Graphics Interchange Format (GIF) for continuous-tone imagery. Use a miniature thumbnail image as a link to the full-size image--the document will load more quickly and readers will fetch only the images they care about. Provide alternative text that describes the image sufficiently and allows the reader to choose whether to load it.

5.4.4 Decoration

Avoid non-functional images.

Carefully selected images can be used decoratively as logos or banners to help create a unified look for a site. Minimize the number of such purely decorative images, and avoid the gratuitous use of images that serve no purpose at all.

5.4.5 Images for Layout

HTML provides only limited capabilities for page layout when compared to typesetting software or word processors. This is deliberate; HTML is designed to be viewed on a variety of devices, ranging from the tiny monochrome screen of a palmtop to a high-resolution true-color monitor. HTML may also be "viewed" by a text-to-speech reader or by an autonomous search engine. Nevertheless, web page authors are tempted to control every aspect of the on-screen presentation, and images are among the tricks used to control layout.

Do not simply make a page one big image.

Avoid at all costs the obvious route of creating an entire page as an image. It may look just right on the screen, but the download time will turn away most visitors and in addition, an image cannot be indexed by search engines and therefore will not be searchable.

Be cautious with invisible spacers and other layout tricks.

Some designers encourage the use of invisible images (transparent images in GIF89a format, or images the same color as the background) to push page elements into the desired location. This trick is harmless if done properly, but it is crucial to use only small images, supply null alternative text attributes (ALT=""), and view the result in a variety of browsers, and with image loading disabled, and at several screen sizes. The presentation will differ in each, and such not-quite-exact layout is typically not worth the effort required to produce it.

Finally, it is possible to create sophisticated layouts by chopping up a larger image and putting the pieces into separate cells of a table. The same rules apply as with the invisible spacer GIFs just discussed: harmless if done well, but easily done poorly.

Design tip: Focus on the information content of your page rather than its appearance. Useful content is timeless; nice appearance depends on current fashion.

5.5 Image Maps

Use image maps when it is a must, and provide an alternate means of navigation.

Image maps allow different parts of a single image to be links to different documents on a server. Image maps are useful, even essential, in many applications. For example, clicking on a geographic map to select a country is a reasonable user interface style. An image map might also be used for a row of navigation icons. Image maps can easily be overused, however, as when the image is simply a picture of text and various phrases are links to different documents. The same result can be achieved using HTML, unless a specific style of text is desired to maintain consistency. Furthermore, image maps are not easily usable by blind visitors or by search engines.

When an image map is appropriate, construct one following the previous recommendations concerning the use of graphics (small file size, useful ALT text, etc.). In addition, provide an alternative means of accessing the links in the image map. A typical alternative is a list or menu of links immediately below the map.

Design tip: If a search engine can't follow links from your image map, it won't be able to index your site.

5.6 Multimedia

Examples of "multimedia" on the web include animations, audio, video, and three-dimensional models. The ability to insert such objects into a web page is very useful if done well. Keep in mind that most multimedia is slow to download and requires special viewers on the client, so not all visitors will choose or be able to enjoy such content.

5.7 Animations

Avoid pointless animations that distract the eye and add little or no value to the web site.

Animations are short sequences of images that play either once only or loop repeatedly, often without sound. Typical formats include animated GIF (GIF89a), MPEG, and QuickTime.

Animated GIFs can be inserted like a normal image in a web page; only the first frame will be shown if the browser cannot perform the animation. GIF89a images are often used in banner advertisements. Animated GIFs are not recommended because the endless looping is typically annoying to the user and because of the large file size required for multiple frames. However, if the concept to be illustrated requires a brief animation, GIF89a is a suitable format.

MPEG and QuickTime formats play either in a separate viewer or directly in the browser window and are better than GIF89a for animations or video clips more than a few frames long. Use them as needed for content. Always warn the user about the file size of the download. Whenever possible provide one or more still frames as preview images to help decide whether to download, or at minimum provide a text description. Recognize that QuickTime works better on Macintosh and MPEG works better on PC and UNIX machines.

5.8 Audio

Never use background audio on a web site.

Most visitors do not want background music unexpectedly blaring when they visit a site. Furthermore, unnecessary audio is a waste of bandwidth. Avoid background audio at all costs.

Short audio clips, and continuous streaming audio, can be used to present audible content if the user is informed beforehand and has the proper software. Much bandwidth is required, so distribute audio judiciously. Describe the content to be provided and the size of the clip if known. Provide a written transcript for archived audio.

5.9 Video

Short video clips, and continuous streaming video, can be used to present visual content if the user has the proper viewing software. Much bandwidth is required, so distribute video judiciously. Describe the content to be provided and the size of the clip if known. Provide a written transcript or still images when possible.

5.10 Three-dimensional Models

3D content, typically in the form of VRML (Virtual Reality Modeling Language) or QTVR (QuickTime Virtual Reality) scenes, is a powerful tool for displaying data and objects not easily viewed in 2D. The technology is still immature, however, and not all users have the software and processing power necessary to render the models, so use such content with caution. Ideally it should only be a supplement to traditional content based on text and images.

Design tip: A video animation that simply shows a solid object moving (for

example, a rotating earth) may be more compactly represented in VRML (as object plus code to make it rotate) than as a movie (as multiple frames to give the appearance of rotation).

5.11 Cross-Browser Compatibility

Ensure that a site is usable in a variety of browsers.

Sites will be visited using browsers both modern and ancient; by Macs, PCs, UNIX boxes, and VMS machines; by people who choose not to load images; by blind users with text-to-speech readers; by power users with big monitors; by wireless warriors with wristwatch browsers and cellular modems; by WebTV watchers on their couches; and especially by search-engine robots. A page that looks great on one screen, with a particular browser, using a specified font size, may be unattractive or unreadable on another screen. It is therefore **essential** to test a page using as many different viewing configurations as possible.

Cross-platform testing can be performed manually, and can be automated in part by third party utilities. After a while it becomes clear what constitutes cross-platform HTML and explicit testing can be done less frequently.
NASA wishes to disseminate information to the broadest possible audience. It is therefore not acceptable to label a site "Best viewed with browser ABC" and omit compatibility testing.

Like all federal agencies, NASA must comply with the Americans with Disabilities Act. Such compliance requires that web sites be accessible to the blind. Either every page must be accessible (readable by a text-to-speech converter, for example), or a parallel set of accessible pages must be provided.

> *Design Tip: If your site is for public consumption then you should test the site on no less than three platforms (Mac, PC, UNIX) using at least two browser brands per platform. A brand of browser such as Netscape or Internet Explorer will display the same site differently on different platforms. Also, be sure to modify the browser settings to turn off image loading for navigation testing.*

5.12 Access Counters

Access counters show visitors the number of other people that have visited a site or page. Counters generally provide information that the user does not care about and that may be embarrassing if there are few hits or if the counter gets reset to zero. Counters add load to the server and may require an extra image in the page. Counters should be considered a vanity feature rather than an element of good page design.

Browser caching or site access through a route other than the home page will cause the counter to sporadically increment. Therefore the counter does not always provide an accurate representation of how often the site is accessed.

However, there may instances such as for reporting purposes when some sort of web server metric is required. An invisible counter is conceptually identical to the typical visible counters that appear on innumerable web pages. These counters simply count the number of times that a web page has been downloaded (this count is different from the number of times that a page is viewed due to local caching considerations). The obvious difference between visible and invisible counters is that for invisible counters there is no number that is physically displayed on the web page. The main reason for using an invisible counter is to avoid clutter on the web page when the number only provides value for the webmaster.

A better approach is for the provider of the web delivery to offer delivery log reports, regularly published on the web (access restricted as necessary), which offer the "big three" web statistics (number of "hits", quantity of data delivered, and unique network addresses served), as well as, individual "hit" scores per web site page. This proactive web usage reporting approach adds value for the user, allowing them to use much broader data in their own internal reporting without cluttering pages.

6.0 Site Accountability

6.1 What should go on the WWW and how is it perceived?

NASA WWW documents are subject to the same accountability and dissemination policies and guidelines as hardcopy documents.

NASA Public Affairs Internet Policy states; "From its origins in 1958 to the present day, NASA has pursued a policy of openness and accessibility. The Agency's business is conducted and explained in full public view, and the results of NASA's activities are considered a resource to be made as widely available as possible. With the advent of the Internet, a medium NASA helped create, a powerful new communications tool is available for explaining NASA's programs to the American people, and for making many of the results of those efforts widely available electronically…NASA employees should follow the same general guidelines for posting information to the Internet that now guide their actions with other media, including television, radio, magazines and newspapers." Documents served from government machines and intended for public consumption are viewed as official government records and publications. Therefore, they are subject to the same legal and regulatory guidelines and policies applied to hard copy materials and publications.

Documents that are in development or are not to be viewed by the public should be placed on WWW servers which protect access to those documents.

All documents, whether public or private shall be directly related to the official responsibilities of individuals and organizations fulfilling their assigned missions and charters consistent with their individual Center, NASA, and the federal government. The World Wide Web Home Pages Guidelines and Best Practices (http://skydive.ncsa.uiuc.edu/cybrary/2/guidelines/), authored by the World Wide Web Federal Consortium (http://hoohoo.ncsa.uiuc.edu:7777/consortium_info/agency_policy_guidelines/agency_policy.html) goes on to further state in section II.B External Linked Content (http://skydive.ncsa.uiuc.edu/cybrary/2/guidelines/content2.html) " The decision to include a link to an external (non-NASA) source should be consistent with sound public policy, in support of the Agency's mission, and based on the WWW site's statement of purpose."

6.2 Policies and Guidelines For Information Dissemination

The Office of Management and Budget (OMB) through it's draft guidelines has clearly indicated that federal Web sites fall under the existing policies governing federal information dissemination. OMB Circular A-130 (http://www.whitehouse.gov/WH/EOP/OMB/html/circulars/a130/a130.html)

The NASA Internet Usage Policy , Executive Notice 02-95 (ftp://ftp.hq.nasa.gov/pub/cio-office/Exc-Notice/en2-95.doc) establishes NASA's position for electronic publication and dissemination of NASA information for public access over Internet. The subject shall be applicable to any information stored on resources in the "nasa.gov" domain.

Government records are subject to a Freedom Of Information Act (FOIA) (http://www.epic.org/open_gov/foia/us_foia_act.html) request. "Those seeking information are no longer required to show a need for information. Instead, the "need to know" standard has been replaced by a "right to know" doctrine. The government now has to justify the need for secrecy." In 1996 Congress amended this act "to provide for public access to information in an electronic format, and for other purposes" through The Electronic Freedom of Information Act Amendments of 1996 (EFOIA) (http://www.epic.org/open_gov/efoia.html).

The World Wide Web Home Pages Guidelines and Best Practices (http://skydive.ncsa.uiuc.edu/cybrary/2/guidelines/) outlines and makes reference to more statues and regulations that should be reviewed prior to disseminating government information. The guidelines document lists these statues in sections I.A Statues and Regulations (http://skydive.ncsa.uiuc.edu/cybrary/2/guidelines/statutes.html) and Appendix B.1, Specific Federal Agency Standards (http://skydive.ncsa.uiuc.edu/cybrary/2/guidelines/agency.html). The guidelines document

continues on to answer more specific questions concerning the Paperwork Reduction Act and OMB Circular A-130 (http://skydive.ncsa.uiuc.edu/cybrary/2/guidelines/legal.html#Paperwork Reduction Act) and how they apply to the Agency's use of the WWW and an explanation of the Federal Records Act (http://skydive.ncsa.uiuc.edu/cybrary/2/guidelines/legal.html#Federal Records Act).

> *Design Tip: The old adage, "better safe than sorry" applies when posting documents to the WWW. If you are not sure if a document should be read by your intended audience, for any reason, then have appropriate personnel review the document before posting it to your site.*

6.3 Should all sites have a NASA person(s) accountable for its existence?

NASA WWW sites should contain the name of the person or organization responsible for the site's content and operation.

In Executive Notice 02-95, Internet Usage Policy (ftp://ftp.hq.nasa.gov/pub/cio-office/Exc-Notice/en2-95.doc), section 1.3.1 states "... there must be clear accountability for the accuracy and appropriateness of the information displayed." Therefore all government sites should display the name of a point of contact capable of explaining what is contained within the site. Missing from this policy statement is how that accountability is to be displayed, should the accountable party be a person or an organization, and what process should be in place so that users who pose questions to the site's accountable party will receive a quality response in a timely manner. Feedback from site visitors is important because it keeps the site's creators in touch with the needs of its audience and provides visitors with a direct way to get information to the site's accountable point of contact. The following paragraphs provide suggestions for completing these tasks.

6.4 Who and what is a site curator and what is their role?

Place the name of the site curator at the bottom of the site's home page at the very minimum and optimally at the bottom of all site pages.

The site curator is responsible for developing, deploying, and maintaining the WWW site. If the site later develops problems such as links that do not work, images that do not display, or the site cannot be found, the site curator will be contacted and is responsible for correcting the problem. The curator can be a civil servant(s) or a contractor(s). It is best practice to adhere to a naming convention to identify the site curator, and maintain it through out the site. Suggested naming conventions are:

Site Curator: John Q. Public (john.q.public@nasa.gov)

Site Construction: John Q. Public (john.q.public@nasa.gov)

Design Tip: It is best practice to include a "mailto: " link to the site's responsible point of contact. You may want to create an email alias (i.e.webteam@nasa.gov) instead of using a specific individual's email account. Email aliases help hide the "email identity" of the site's point-of-contact (POC) and allows for the addition or removal of multiple POCs with modification to the site's mailto: link. Whether using an email alias or a specific name, the full electronic mail address should also be displayed for those visitors who cannot or have not configured the electronic mail capability within their browser application.

6.5 Who and what is the NASA Official Responsible For Content and what is their role?

Place the name of the NASA Official at the bottom of the site's home page at the very minimum and optimally at the bottom of all site pages.

Site officials are responsible for the content of the site. They are responsible for approving all information to be posted to the site and reviewing that information once it has been posted to the site. The use of a staging server where pages may be approved before publishing may assist in this process. The official is also responsible for periodically reviewing the site's content and ensuring the site is kept relevant, accurate, and up to date. The NASA Official must be a civil servant.

As with the site curator it is best practice to adhere to a naming convention to identify the NASA Official, and maintain it through out the site. Suggested naming conventions are:

NASA Official Responsible For Content: John Q. Public (john.q.public@nasa.gov)

Page Content: John Q. Public (john.q.public@nasa.gov)

NASA Point Of Contact: John Q. Public (john.q.public@nasa.gov)

Design Tip: It is best practice to include a "mailto: " link to the site's responsible point of contact. You may want to create an email alias (i.e.webteam@nasa.gov) instead of using a specific individual's email account. Email aliases help hide the "email identity" of the site's POC and allows for the addition or removal of multiple POCs with modification to the site's mailto: link. Whether using an email alias or a specific name, the full electronic mail address should also be displayed for those visitors who cannot or have not configured the electronic mail capability within their browser application. In addition, adding a default subject line to the mailto: will allow the site curator receiving the email to easily determine which Web page is referenced, which the visitor may not include in their email.

6.6 Timely response to site inquiries

All site inquiries should be answered in a timely and accurate manner, and to the best of a person's or organization's ability.

The Yale Style Guide (http://info.med.yale.edu/caim/manual/contents.html) states "Feedback also means being prepared to respond to your user's inquiries and comments. ... Planning for this kind of ongoing relationship with the users of your site is vital to the long-term success of the enterprise."
The most personal approach is to respond to each electronic mail inquiry with a personal response that, at the very least, acknowledges receipt of each inquiry. However, this can become a very time consuming task. Therefore the next logical step is to guide users to ask questions through a feedback form. Use the feedback form to

⟨ Help filter inquiries. Use radio buttons to ask the user if they expect a response. Often visitors do not, they just want to give the site curator comments.

⟨ Provide an immediate response to all visitors through a "canned email" that is sent by the script processing the form's information.

⟨ Provide visitors the capability to select their subject of interest. Based on the selection send electronic mail directly to personnel within an organization who will provide the highest quality answer in a timely manner. This distributes the responsibility of responding to inquiries to many people and organizations rather than placing the burden on one person.

⟨ Try to redirect the visitor to a Frequently Asked Questions site. They may have a common question that has been asked many times before.

Design tip: The person on the other side of the mail message is another human being. They don't know you have been asked the same question ten times this week, or what is obvious to you apparently is not to them. Your response is a reflection on NASA and at all times should be clear, concise and civil.

7.0 Site Security

NASA has become very dependent upon the WWW to disseminate information both internally to its employees and externally to the general public. To maintain the integrity of NASA's information and systems, this information needs to be adequately protected, yet readily available to all intended audiences.

7.1 What is the current NASA policy on protecting WWW pages?

Determine the security requirements and place the information only on WWW servers that are configured to meet required security controls and with personnel to maintain those security controls.

Procedures and guidelines for determining the appropriate measures necessary for protecting information and for determining protective controls required for computer systems which host WEB pages can be found in the NASA Procedures and Guideline for the Security of information Technology (NPG 2810).

Also, you may want to review the NASA Firewall Strategy, Architecture, Standards and Products, NASA-STD –2813 (http://www.hq.nasa.gov/office/cio/standards/2813.pdf) (pdf, 102K) details the firewall architecture adopted by the Agency. The architecture is based on the following three information service "islands":

1. <u>Center Private Network:</u> Center information technology resources that require protection from the global Internet. These are often referred to as NASA-only or Center-only (with the Center's name being substituted for "Center") sites.

2. <u>Center Public Network:</u> Center information technology resources that must be accessible by the public-at-large, but at the same time, requires very substantial protection to assure uncompromised information integrity and service availability. Public WWW sites are included in this group

3. <u>Center Open Network:</u> Center IT resources which need to be freely accessible, via the Internet, by the scientific community and the public-at-large, and require protection in terms of data integrity and availability, but without firewalls.

> *Design Tip: Test the security of Center-only or NASA-only through access from your or a friend's home account. You should receive an "access denied error " (403) when you attempt to display a site blocked to the general public.*

7.2 Who makes the determination of what information should be secured?

The "data owner" is responsible for determining the dissemination requirements for their information. These dissemination requirements drive the security controls required for the information and the WWW sites.

The author of the document must make the first level of determination. When the data owner has determined the dissemination requirements for the information contained on a web site, the data owner must then place the document on a WWW server that has security controls in place that can ensure the appropriate dissemination requirements.

Design tip: Do not always assume that all documents on a site have the same security requirements. Site document composition varies and sometimes portions of a site must be maintained more securely than others. Evaluate each document independently.

7.3 What are the available security controls?

WWW data owners must make the information's dissemination requirements clear to WWW server administrators,. WWW server administrators must implement and maintain security controls required by WWW document authors. Some examples of security controls are:

Machine or Site Authentication: Offers protection for the entire site, individual documents or directories by allowing only specified IP addresses, IP subnets, or domains to access them.

End User Authentication: Users must enter a name and password before access is granted.

Encryption: The document request and the document are both encrypted (altered) such that the document cannot be read by anyone other than the intended recipient. This is also reliable for end user authentication.

7.3.1 How secure are these controls ... really?

Machine or Site Authentication: Hackers can "spoof" IP addresses and make it seem as though they are coming from an IP or domain other than their real one. This will allow them to bypass this authentication type.

End User Authentication: A poor password can be easily guessed. Passwords are also transmitted from browser to server "in the clear" each time a document is requested from the server, thus, leaving passwords open to interception by hackers.

Encryption: This is the most secure method for transmitting documents over the WWW. This method requires the secure exchange of either a single key (symmetric cryptography) or the secure publication and storage of public and private keys (asymmetric cryptography).

7.3.2 How Do I Secure a Web Server?

The World Wide Web Consortium Security Resources site (http://www.w3.org/Security/) covers in detail how to run a secure server and how to

protect documents.

The WWW Federal Consortium Guidelines document contains a WWW security checklist
(http://hoohoo.ncsa.uiuc.edu:7777/consortium_info/agency_policy_guidelines/agency_policy.html).

7.4 What security technology is maturing and is NASA doing anything with it? (e.g., Entrust Certificates)

Ames Research Center (ARC) is working with other NASA Centers to develop a NASA Public Key Infrastructure (PKI) policy. The PKI project is aimed at designing and constructing an infrastructure where applications can utilize this infrastructure to incorporate security features. One such security application that ARC has demonstrated is secure e-mail.   The system will include all necessary components including a Certificate Authority, a Center Registration Authority, and X.500 Directory Service for certificate repository.   This system will provide the full certificate life cycle management for certificate issuance, key recovery, key history, and key management.

7.5 What if a site has a security breach?

Immediately report WWW security breaches to your Center IT Security Office. Do not take immediate corrective action as evidence of the breach may be destroyed.
Ames Research Center (ARC) is the Principal Center for Information Technology Security (http://george.arc.nasa.gov/ITS/). ARC maintains a list of the Center Information Technology Security Offices
(http://george.arc.nasa.gov/ITS/principal_center.html#anchor586513) and outlines the steps for reporting suspected security breaches
(http://george.arc.nasa.gov/ITS/principal_center.html#anchor583927).
NASA Automated System Incident Response Capability, NASIRC (http://www-nasirc.nasa.gov/nasa/index.html) maintains a checklist and a FAQs site to help secure a system and answer operating system specific questions.

> *Design tip: Use products that automate monitoring of WWW servers and their underlying operating system and provide automated notification should a security breach occur.*

8.0 Accessibility

Site audiences need to be defined in order to determine access boundaries. Once those boundaries are set they should be enforced and should be declared as part of site identification or registration with a Center's site registration authority. Sites fall into three access categories: (1) public (open to each and every member of the public with an internet browser--this is NOT to say all versions of every browser); (2) limited (either by a broad group of users, an organization, or a Center and enforced by addressing limitations); and (3) restricted (limited to a special kernel of users who supply X.509 or username/password authentication before any site access is permitted.

Sites should only be accessible to all persons when the site is declared with Center registrars to be "public" (using definition above) in nature. Sites of either limited or restricted classifications shall NOT be available to all users. While a limited site may be visible to the public, only members of that site's declared audiencemay access the system.

Public sites should be on public networks. Other sites should be on NASA-specific networks.

9.0 Web Marketing

Due to the sheer numbers of web pages on the Internet, one of the first steps after developing a web site is to ensure that potential visitors can find it. Whether an organization is large or small, in order to be found on the Internet, the first step of web site marketing is to get listed on all of the major search engines and directories. Most people turn to the search engines and directories to start their search for a specific product or service. By establishing a presence there, the chance of drawing visitors greatly increases. Getting listed in directories and search engines is the most logical and cost-effective means to placing a site where potential visitors start their searches on the Internet.

The following information is designed to help promote a web site in the most popular search engines on the Internet. Most of the following information was obtained from the "Search Engine Secrets" topic on the PegasoWeb site at

http://www.pegasoweb.com/freerep.htm

9.1 Overview of Search Engines

Generally speaking search engines read a page and memorize words on the page that are referenced when a user (searcher) performs a search. In simple terms if the word that the 'searcher' enters into his query is not on the page there will be no match and that page will not be listed. However each search engine uses a different technique for locating words on a page. Some search engines look for keywords or "meta-tags" that are actually "behind" a

page and are not visible to the person viewing it. Search engines generate a Confidence or Relevance Factor. Sometimes they display this as a percentage or points. The Relevance Factor is determined not only by how many times the words used in a query are in a document, but also how close they are together.

9.2 Choosing Keywords

Choosing keywords can be fun! It is an art. Be as specific as possible. However, if unusual words or words that are too unique are used then the 'searcher' will probably not include them in his or her search query. Try to consider what the 'searcher would do. Make a list of the keywords that cover the information best. Make certain to include these words whenever possible in the page.

The more frequently these words are included in the document, the higher the Relevancy Factor, and the higher up the page will be listed in the search engines' list. However, do not overuse them to the point that it will take away from the readability of the document. Some search engines will actually reject the keyword if it is repeated too often or spammed in the document. Others will reject the entire document if the keywords are not mentioned in the body of the document.

9.3 The First Change to Make is The Title

One of the first things to look at and change if necessary is the page title. The page title is not what is on top of the page, but is what is displayed at the very top of the browser window. Looking at the HTML code, it is what is in between <TITLE></TITLE>. Many search engines start with the title. If keywords are not contained within the title, then the Relevance Factor can go down.

Use some of the most relevant keywords in the title. The title should catch the searcher's eye, but it should also indicate to the searcher what is in the page. When a searcher performs a search, the search engines will display the title of the page, and sometimes the first part of the page as a description. If the title doesn't indicate what is in the page a searcher may not bother to visit the page.

9.4 How To Use Meta-Tags

 Meta-Tags are placed "behind" the page. In addition to the contents of a page, several search engines rely on the content of meta-tags to get the description and keywords for a web page. The following is taken from one of the most popular search engines used today. -

Specifying an exact description for to use:

Create a description for the Web page, rather than having to use the first 200 characters. Use the<META> tag within the <HEAD> element to do this. The basic syntax is:

<META name="description" content="Write your description here">
Do not use any HTML tags within the description or content part of the META tag.
Suppose the HTML at the top of the Web page looked like:
<HEAD>
<TITLE>The Kids In The Hall Home Page</TITLE>
<META name="description" content="Home page for The Kids In The Hall, Canada's funniest comedy group.">
</HEAD>

The following title and description would appear when the page is listed:
The Kids In The Hall Home Page
Home page for The Kids In The Hall, Canada's funniest comedy group.
Specifying keyword phrases:
A second <META> tag can be used to specify keyword phrases that further describe the Web page. The basic syntax is:
<META name="keywords" content="Write your keywords here, in a comma separated list">
For example, if the HTML at the top of the page looked like:
<HEAD>
<TITLE>Landing Venture Capital</TITLE>
<META name="description" content="Tips on obtaining venture capital for your hi-tech start-up.">
<META name="keywords" content="asset based lending, capital, enterprise funding, equity funding, expansion capital, funds for growth, growth financing, underwriting, venture investment">
</HEAD>

The search engine will take it into account when it indexes the page. Using the <META> tag to add keywords to a page simply provides extra information about a page without that information appearing to the reader.

Please note that the search engine indexes the entire page (except any text within comments), regardless of whether or not a description or keywords are included in the <META> tags. The words in the <META> tags are indexed in addition to the rest of the document.

Some search engines use meta-tags to get the description of the page and then list this on their index. Construct the description carefully. It should get the searcher's attention, as well as describe the page. It is also good to use several of the top keywords in the

description. Some search engines limit the description to 25 words or less. Keep it concise.

9.5 When in Doubt Copy Others

Perform a search for pages with similar content using multiple search engines such as http://www.altavista.com, http://www.lycos.com or http://www.infoseek.com. See which pages show up at the top of the indices. Go to those pages and note how each has used the keywords used in the search. Also check the document source to see if meta-tags have been used for keywords and description. Doing this on a few pages should provide some good keywords to add.

9.6 Testing Keywords

Similar to the technique above, keywords can be tested. A developer should try chosen keywords, and also combinations of chosen keywords on several search engines. See what types of pages are returned. This will give an idea whether the chosen keywords are good choices or not. If certain combinations of keywords seem to target the appropriate type of pages better than others, then moving these keywords to the beginning of the page might improve the page's listing in the search engine.

9.7 The "Invisible" Keywords Secret

The following is a technique, which when used properly, can dramatically increase the relevancy percentage used by search engines to put a page at the top of their list. Go to the bottom of a web page. Perform a few carriage returns until the last line of text is no longer visible. Using the smallest type size on the browser, type the list of keywords. Change the font color of the keywords to the same color as the background. This will make them invisible to the reader but not to the search engines.

Be warned that in order to prevent abuse, some search engines have no tolerance for stuffing a web page and will subsequently reject any further requests to add a stuffed page to their search engine.

9.8 Got Images? Use Them to Hide Keywords

Images can be a great place to hide keywords in a document. When inserting an image into a document, there is an option to add text that will show if the image isn't displayed. In some HTML authoring programs this is sometimes called an alternate.

The actual HTML code is as follows:

<IMG SCR="yourgraphic.gif" HEIGHT=30 WIDTH=60 ALT="describe your graphic and then list your keywords">

Once again, consider the warning above in section 9.7 about stuffing a page.

9.10 Hide Keywords in Comments

HTML programmers use comments to add remarks about the code. Comments are invisible and sit behind the page. They can only be read by reading the source code. But search engines can read them. A comment is placed in between < >. For example

<rich money prosperity gold wealth>.

Once again, consider the warning above in section 9.7 about stuffing a page.

9.11 Getting Listed On Top Even If Someone Is Using The Same Keywords

Some search engines will alphabetize their lists. If another site uses the same keywords one way to get to the top is to start the title with an "A". For example, if the title was "Wonderful Books", changing it to "Absolutely Wonderful Books" may improve the site's standing. Numbers are also good. In fact sometimes numbers are indexed ahead of the alphabet. It is best if the "A" is part of a word, not just the letter "A". Some search engines do not consider a single character when searching.

9.12 Use S's For Super Search Results

Put an "s" at the end of keywords to make them plural. For example, it the keywords "dog lover" were used and a user queried a search engine with the word "lovers", the page would not show up on the list. If the keyword "lovers" was used, search engines will be more likely to detect it whether the query word used for the search was "lover" or "lovers." "Lover" is part of the word "lovers" and most search engines will pick this in a search.

9.13 Using The Most Searched Words

Find out and keep abreast of what words are searched for the most and try to include some of them as keywords. For example, by using the title "Dog Training For Men and Women" instead of just "Dog Training" the title would include what is currently the fifth most used search word on the internet.

9.14 Submit Multiple Pages

An old trick is to submit multiple pages to a search engine. For example, identical pages might be named

http://yourmoney.com/
http://yourmoney.com/money.htm
http://yourmoney.com/wealth.htm

This is also good exercise in seeing how each search engine ranks a page depending on the differences in the title.

9.15 Where To Submit A Web Page

The first place to submit a web page is to the major search engines. The following is a list of some these engines.

Altavista Search Engines
Error! Bookmark not defined.

Webcrawler Search Engines
Error! Bookmark not defined.

Lycos Search Engines
Error! Bookmark not defined.

Infoseek Search Engines
Error! Bookmark not defined.

Excite Search Engines
Error! Bookmark not defined.

HotBot Search Engines
Error! Bookmark not defined.

Magellan Search Engines
Error! Bookmark not defined.

Yahoo Search Engines
Error! Bookmark not defined.

When submitting a site, be patient. Due to volume and backlog some search engines can take over 6 weeks to get a site listed. Others will show up in a couple of days. Do not be discouraged if a page does not show up immediately.
There are many "links" pages that will list sites. Perform a search to find these sites. There are directories, yellow pages, and other "links" pages that will list a site for free. Find sites that would be possible candidates for "reciprocal" links.

Find newsgroups that may have readers interested in specific sites. Post the site's address to the newsgroup with a catchy heading. Similar to newsgroups, sometimes there are mailing lists that may have subscribers that would be interested in specific sites. Subscribe to the list and post a message.

9.16 Where To Test Keywords And Position Ranking

Monitor the rank of a Web site among the top 12 search engines for selected keywords to help improve listings and get more hits.

PositionAgent
http://www.positionagent.com

9.17 The Way To Submit to 100's of Search Engines and Directories

Announcement Services provide a way to simultaneously submit a site to many search engines and sites that will list a web site. Many of these are free. The following is a list of some of these Announcement Services. At the time this list was made they were free.

Worldwide Broadcaster - This is a relatively new announcement service that allows submission to over 200 sites and search engines with only one submittal to Broadcaster!
http://www.broadcaster.co.uk/

Postmaster - Posts a site to 24 different places and search engines for free.
http://www.netcreations.com/postmaster/

Submit It! - Submits a site to 15 places and search engines for free.
http://free.submit-it.com/

Quick Launch - Submits a site to 15 directories and search engines for free.
http://www.qwiklaunch.com/

10.0 Site Search

Within a local search page a good practice would be to add an input panel or link to the Agency-wide search engine so that a visitor could look beyond a local site if needed. If appropriate, also include a link to the NASA Image Exchange (NIX) image search engine. The goal is to accommodate the user and provide the broadest possible set of tools to aid in searching for information.

Since it will not be known at what point a user may enter a site (either at the top or to sub-level pages), it is recommended that a link to the site search page be provided on every page of a site to assist with further inquiry within the site.

The Webmaster should encourage web content providers who desire their data to be indexed in a meaningful way by a search engine to make heading descriptive! This can be accomplished by including important keywords in the heading and using HTML heading tags (e.g. H1, H2, H3) rather than just adjusting the font using <FONT SIZE="x"> HTML tags. Most robots index the FULL content of HTML heading tags!!

It is also helpful to provide a link from the search page to a help page with simple instructions giving the syntax and examples for a particular search engine.

Selecting a search engine, which also searches against metadata keywords, can help direct users to the most appropriate materials. The use of metadata keywords within a site will also accomplish the same for many Internet-wide search engines and assist in promoting a web site.

Selecting a search engine which uses metadata descriptions in the results summary will allow control of the words displayed and aid a user on which link to follow. The use of metadata descriptions within a site will also accomplish the same for many Internet-wide search engines and assist in promoting a web site.

Make certain that the update frequency of a search engine index is appropriate for the frequency of updates within that site. Users can become very frustrated if a search returns 404 errors for pages that no longer exist.

Use appropriate wording on pages so that the search will find items. For example, never use an acronym on a page without spelling it out at least once. Consider the user who may not be familiar with certain projects or programs.

Test to make sure that search results for both simple and complex searches are fast enough to avoid frustrating the searcher. If search returns are too slow, consider fewer results per page, such as 10, so that less HTML code is sent.

Consider the following features when choosing a site search mechanism: results sorted by relevance, natural language queries, search operators for greater control, secondary searches within the current results, find similar documents, use of metadata keywords, and use of metadata descriptions.

Webmasters that want to limit robot access to sensitive or private data on Webservers should employ the "robot.txt" file. A Webmaster can easily restrict what directories and files a robot can access (HTML or otherwise) just by adding simple control parameters into the "robot.txt" file.

11.0 Authoring Tools

Because software versions are regularly updated, and because NASA cannot appear to endorse commercial products (even those purchased) this will not be a discussion of the pros and cons of particular applications. Rather, these are a few general principles to keep in mind when evaluating and purchasing software for editing Web documents.
As a center Headquarters has baselined software packages
(http://www.hq.nasa.gov/help/hqwsstd.htm), including HTML and graphics editors.
Most Center information technology security groups discourage downloading untested software, even demos from software publishers' reliable sites. When testing a product, it is best to coordinate with the Center IT group to obtain demonstration and test software.

11.1 Hypertext Markup Language (HTML)

Hypertext Markup Language is a markup language in which embedded tags are read by a Web browser and converted into a Web page. The language has evolved since its invention in 1994, incorporating as standard tags many that were developed for use within specific browsers.

Use an HTML editor that recognizes the latest version and one version older of HTML, as defined by the World Wide Web consortium. As of mid-1998, that would be HTML 3.2. HTML 4.0 has been released, but many browsers do not yet fully support it. This is likely to continue to be the case with each new release.

HTML editors have blossomed from a few small applications that allowed only for hand coding to more complex software with a graphical interface (also known as "what you see is what you get" or WYSIWYG). The latter type can make page creation much simpler, but within the graphical interface there may be programming that effectively limits some options. For example, inserting graphics into a Web page via "drag and drop" may generate relative Hypertext links that will have to be corrected before posting the page to a Web server.

Some programs will allow switching back and forth between the graphical interface and HTML code, which can help overcome some of these limitations.

If considering "WYSIWYG" editors, choose one that still allows editing of the HTML source code. Preference should be given to those that do not try to "second guess" what was intended, but leaves code as is.

Learn HTML. If a Web page does not appear as intended, "WYSIWYG" editors are no substitute for knowing HTML code.

The following are resources (from "World Wide Web Home Page Guidelines and Best Practices", prepared by the World Wide Web Federal Consortium, revised November 1996; scheduled for revision mid-1998.):

Putting Information onto the Web
(http://www.w3.org/hypertext/WWW/Provider/Overview.html#author)
a collection of documents on authoring hypertext

NSF/NCSA World Wide Web Federal Consortium Training Materials Page
(http://hoohoo.ncsa.uiuc.edu:7777/consortium_info/agency_policy_guidelines/agency_policy.html)
Style Guide for On-line Hypertext (http://www.w3.org/Provider/Style/Overview.html)
Tim Berners-Lee, World Wide Web Consortium

NCSA HTML Style Guide (http://www.ncsa.uiuc.edu/styleguide)
National Center for Supercomputing Applications

Introduction to HTML and URLs
(http://www.utoronto.ca/webdocs/HTMLdocs/NewHTML/intro.html)
Ian Graham, University of Toronto

Guide to Writing HTML Documents
(http://www.hypernews.org/HyperNews/get/www/html/guides.html)
Dan LaLiberte, National Center for Supercomputing Applications

Composing Good HTML (http://www.cs.cmu.edu/~tilt/cgh/)
Eric Tilton

Yale C/AIM WWW Style Manual (http://info.med.yale.edu/caim/manual/contents.html)
Patrick J. Lynch, Yale Center for Advanced Instructional Media

Yahoo collection of WWW-related sites and documents
(**Error! Bookmark not defined.**)

Names and Addresses, URIs, URLs, URNs, URCs (http://www.w3.org/Addressing/)

11.2 Graphics

There are a variety of formats available for graphics that can be incorporated into a Web page. The two most commonly used are JPEG and GIF. Each allows for compression of images, using different techniques.

The important point in editing graphics for the Web is to choose hardware and software that give the most flexibility on image resolution. Resolution is usually expressed as pixels-per-inch or dots-per-inch (dpi). The higher the resolution, the larger the graphics file and the longer the download time.

A standard computer screen displays images at 72 dots per inch (dpi); higher resolutions are unnecessary for graphics meant to appear only on a monitor. Laser printers generally require higher resolution (120 dpi and above) and files meant for use in professional printing begin at 300 dpi and rise quickly. Any worthwhile graphics program will have the capability to change an image resolution, but it's best to start with a high-resolution graphic and reduce the resolution. Going from high to low resolution generally gives better results than going from low to high.

If high-resolution graphics need to be part of Web site, they should be clearly marked as large files. Small versions, called "thumbnails" should be posted to allow users to browse and select without having to download all the files. (See "Page Design" and "Site Design".)

Use graphics hardware and software that at minimum will create graphics in JPEG and GIF formats, which are viewable in both major commercial Web browsers and across platforms.

Note that Web browsers may display fewer colors than are available in graphics applications. A graphic created with a 256-color graphics palette may appear slightly different in a Web browser with a 216-color palette.

Use software that includes a Web color palette, which displays graphics as they will appear on the Web.

11.3 Plaintext Files

Plain-text files appear in Web browser windows exactly as described: plain, with no style, font or size variations. They can be useful for making large amounts of data accessible, but are more useful when transferred via the File Transfer Protocol (FTP) than when browsed on the Web. In some cases, such as NASA HQ Press Releases, plain-text is a

legacy from the early days of the Web, when text files were easier to create than HTML and Webmasters still had concerns about the download times of even small HTML files.

Limit the amount of plain-text files on a Web site, using HTML to take advantage of the Web's graphic capabilities.

Any word-processing program should be capable of saving documents in a "text only" format with carriage returns (distinct from paragraph markers) at the end of each line. Before converting a formatted document using a save function, convert the font to a fixed-width font (e.g., Courier) and adjust margins and spacing manually to get a better idea of how the document will look. This will make for fewer adjustments after the conversion. Bear in mind that most Web users will choose their own fonts and sizes, so a document may appear differently to different users.

If the original graphic appearance of a file is essential in order to convey its information, use a format, such as Adobe's Portable Document Format, to maintain the original look and feel. These kinds of programs are also useful for conveying material, such as reports, that users may want to print and read off-line. If for some reason this option is not available, create a plaintext file. Use a commonly available fixed-width font with margins that fit completely within the window of a standard 13-inch monitor. Maximum character width should be about 60 characters. Include a note at the top of the file indicating the font, size and margin width used in original formatting.

Always double check documents within a standard Web browser before posting. Some programs embed formatting commands that are not always deleted in conversions to text (e.g., "smart quotes" in Microsoft Word). These formatting commands then appear as indecipherable characters (Õ) in the browser.

11.4 Multimedia

The types of multimedia programs (sound, video, animation) available are too numerous to discuss in great detail. When incorporating multimedia into a site, bear in mind considerations of download time, audience interest, availability across platforms and accessibility requirements.

Use multimedia software that creates files viewable in the two major commercial Web browsers without requiring an Internet user to purchase additional software. (Established software that requires users to download free "plug-ins" or "players" from reliable sites is acceptable as long us such plug-ins are available across platforms.)

Appendix A: Acroynm List

Acronym	Description
CIO	Chief Information Officer
DPI	Dots per Inch
EFOIA	Electronic FOIA
FAQ	Frequently Asked Questions
FOIA	Freedom of Information Act
FTP	File Transfer Protocol
GIF	Graphics Interchange Format
HQ	Headquarters
HTML	HyperText Markup Language
IP	Internet Protocol
IT	Information Technology
MPEG	
NASA	National Aeronautics and Space Administration
NASIRC	NASA Automated System Incident Response Capability
NHB	NASA Handbook
NIITA	NASA Integrated Information Technology Architecture
NMI	NASA Management Instruction
NPD	NASA Policy Directive
NPG	?NASA Procedures and Guidelines
OMB	Office of Management and Budget
PC	Personal Computer
PDF	Portable Document Format
PKI	Public Key Infrastructure
POC	Point of Contact
QTVR	QuickTime Virtual Reality
SSI	Server Side Includes
SSL	Secure Socket Layer
TIFF	
URL	Universal Resource Locator
VRML	Virtual Reality Modeling Language
WBP	Web Best Practices
WWW	World Wide Web
WYSIWYG	What You See Is What You Get
XML	eXtensible Markup Language

Appendix B: Use of the NASA Insignia and Seal on Agency Web Sites

Webmasters should be careful in using the NASA Seal, the NASA Insignia (known to employees as "the meatball") and the NASA Logotype (a.k.a, "the worm") on their Web pages. Proper use of these "visual identifiers" is literally a federal case, outlined in the Code of Federal Regulations (14 CFR §1221.1). The Code charges the NASA Graphics Coordinator in the Office of Public Affairs with ensuring the proper use of the NASA emblems in any setting in which they appear. At the time this document was written, the NASA Graphics Coordinator was Mr. Bert Ulrich, Public Services Division, Office of Public Affairs (phone: 202/358-1750).

This guidance deals with use of the NASA Insignia. The NASA Seal is reserved for use in association with the NASA Administrator. The NASA Logotype has been retired, and should only be used in historical context or with the approval of the NASA Graphics Coordinator.

The Public Services Division has developed standards for using the NASA Insignia on the Web. A Web site containing the guidelines and templates should be available in Fall 1998. For the time being, a limited number of CD-ROMs containing templates and usage guidelines are available from the NASA Graphics Coordinator.

A few guidelines apply to use of the NASA Insignia:

• It must never be used in such a way as to imply the Agency's endorsement of another party's goods or services.

• It should not appear on non-NASA Web pages, and NASA employees should not give permission for other groups to use the Insignia. This prohibition extends to contractors working directly on NASA projects and to organizational "partners" in joint projects. While some groups have appropriated the Insignia for their Web pages, these uses have not been approved by the NASA Graphics Coordinator or the Office of the General Counsel. To resolve questions, contact Mr. Ulrich at 202/358-1750.

• It should never be used on NASA Web sites as a "hot link" except to the NASA Home Page: **Error! Bookmark not defined.**.

Appendix C: NASA Privacy Statement

This policy establishes how NASA will use information we gather about you from your visit to our website. The privacy of our customers is of utmost importance to NASA. If you visit a NASA site...

⟨ **To read or download information:**

We may collect and store information for statistical purposes. For example, we may count the number of visitors to the different sections of our site to help us make them more useful to visitors. Similar information is gathered for anonymous ftp, remote account login, or for other comparable types of connections.

⟨ **To send us an E-mail:**

By sending us an electronic mail message, you may be sending us personal information (e.g., name, address, E-mail address), as in an official Freedom of Information Act request. We may store the name and address of the requester in order to respond to the request or to otherwise resolve the subject matter of your E-mail.

⟨ **To register:**

Some of our sites ask visitors who request specific information to fill out a registration form. For example, vendors looking for marketing opportunities by visiting our Electronic Grants or NASA Acquisition Internet Service sites may be asked to "register" to obtain copies of Requests for Proposals or other NASA opportunities. Other information which may be collected at these sites through questionnaires, feedback forms, or other means, enable us to determine a visitor's interests, with the goal of providing better service to our customers.

We want to be very clear: regardless of the information being transmitted to NASA, we will protect all such information consistent with applicable law.